Journal of Experimental Psychology: Learning, Memory, and Cognition

A Further Examination of Word Frequency and Age-of-Acquisition Effects in English Lexical Decision Task Performance: The Role of Frequency Trajectory

Barbara J. Juhasz, Melvin J. Yap, Akila Raoul, and Micaela Kaye

Online First Publication, April 23, 2018. http://dx.doi.org/10.1037/xlm0000564

American Psychological Association

Journal of Experimental Psychology: Learning, Memory, and Cognition

0278-7393/18/$12.00 http://dx.doi.org/10.1037/xlm0000564

A Further Examination of Word Frequency and Age-of-Acquisition Effects in English Lexical Decision Task Performance: The Role of Frequency Trajectory

Barbara J. Juhasz
Wesleyan University

Melvin J. Yap
National University of Singapore

Akila Raoul and Micaela Kaye
Wesleyan University

Word frequency is an important predictor of lexical-decision task performance. The current study further examined the role of this variable by exploring the influence of frequency trajectory. Frequency trajectory is measured by how often a word occurs in childhood relative to adulthood. Past research on the role of this variable in word recognition has produced equivocal results. In the current study, words were selected based on their frequencies in Grade 1 (child frequency) and Grade 13 (college frequency). In Experiment 1, four frequency trajectory conditions were factorially examined in a lexical-decision task with English words: high-to-high (*world*), high-to-low (*uncle*), low-to-high (*brain*) and low-to-low (*opera*). an interaction between Grade 1 and college frequency demonstrated that words in the low-to-high condition were processed significantly faster and more accurately than words in the low-to-low condition, whereas the high-to-high and high-to-low conditions did not differ significantly. In Experiment 2, an advantage for words with an increasing frequency trajectory was also supported in regression analyses on both lexical decision and naming times for 3,039 items selected from the English Lexicon Project (Balota et al., 2007). This was replicated in Experiment 3, based on a regression analysis of 2,680 words from the British Lexicon Project (BLP; Keuleers, Lacey, Rastle, & Brysbaert, 2012). In all analyses, rated age-of-acquisition also significantly impacted word recognition. Together, the results suggest that the age at which a word is initially learned as well as its frequency trajectory across childhood impact performance in the lexical-decision task.

Keywords: word frequency, frequency trajectory, lexical decision, word naming, age-of-acquisition

The average 20-year old English speaker's vocabulary consists of approximately 42,000 individual lemmas (range 27,100 to 51,700; Brysbaert, Stevens, Mandera, & Keuleers, 2016). This impressive vocabulary is acquired over the course of childhood, adolescence, and young adulthood, with the first word being produced typically before a child's first birthday. The current study examines the extent to which a reader's experiences with words influence skilled word recognition. Specifically, we consider the impact of frequency trajectory (the frequency with which is a word is encountered during reading development) and age-of-acquisition (AoA; the age at which the word is first learned) on lexical decision (i.e., classifying letter strings as words or nonwords) and speeded pronunciation (i.e., reading words aloud) performance.

Word Frequency, Age of Acquisition, and Frequency Trajectory

Words that are encountered more often are processed faster across a variety of tasks (e.g., lexical decision, word naming, and reading words in sentences). Examinations of large-scale databases of lexical decision times (LDTs) have demonstrated that a word's frequency of occurrence is the best predictor of LDTs (Brysbaert et al., 2011). However, how frequent a word is in a given language can be measured in a variety of ways. For English, various word frequency corpora count how often a given word occurs in written texts (e.g., Kučera & Francis, 1967; CELEX Lexical Database: Baayen, Piepenbrock, & Van Rijn, 1995; Zeno, Ivens, Hillard, & Duvvuri, 1995), in Internet discussion groups (HAL; Burgess & Livesay, 1998), or more recently from transcribed subtitles (SUBTLEX; Brysbaert & New, 2009). Studies that have directly compared various word frequency measures have demonstrated that frequency corpora vary in their utility (e.g., Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Brysbaert & New, 2009).

Barbara J. Juhasz, Department of Psychology, Wesleyan University; Melvin J. Yap, Department of Psychology, National University of Singapore; Akila Raoul and Micaela Kaye, Department of Psychology, Wesleyan University.

This research was supported by a Wesleyan University project grant. We thank the undergraduate research assistants who contributed to this project. Portions of the data were presented at the Psychonomic Society Annual Meeting in November 2015.

Correspondence concerning this article should be addressed to Barbara J. Juhasz, Department of Psychology, Wesleyan University, Judd Hall 508, Middletown, CT 06459. E-mail: bjuhasz@wesleyan.edu

Thus, it is important to assess the size of the corpus, where it is compiled from, as well as its age when determining the most valid measure of word frequency to use to assess individuals' experience with words (Brysbaert et al., 2011). Given the ubiquity of frequency effects in lexical processing, all models of word recognition incorporate some mechanism to account for these effects (for a discussion, see Gomez, 2012). In addition, there is also debate about how best to model the form of the frequency effect that is observed in lexical decision (Murray & Forster, 2004), and whether contextual diversity (the number of contexts or documents a given word appears in) is a more important variable than a simple count of how often a word occurs in text (Adelman, Brown, & Quesada, 2006).

Another variable that has been found to reliably predict lexical processing in a variety of tasks is rated AoA. Words that are rated as being acquired earlier in life are typically recognized and responded to faster than those that are rated as being acquired later in life. A recent study by Brysbaert and Cortese (2011) demonstrated that rated AoA was a significant predictor of naming and lexical decision performance, even after controlling for optimal measures of objective word frequency. That being said, despite wide-ranging support for the influence of AoA (for reviews, see Johnston & Barry, 2006; Juhasz, 2005), there is still controversy surrounding this variable in the word recognition literature. Some of this controversy pertains to how AoA is measured. The fact that AoA is often rated by adults (typically college students) has been criticized by researchers who note that adults will not be accurate at determining the exact age or order with which they acquire words and therefore may rely on other property of the words (such as their frequency, ease of creating a mental image, emotionality) to make their rating (e.g., Baayen, Milin, & Ramscar, 2016; Bonin, Barry, Méot, & Chalard, 2004). In fact, children (8–10 years of age) have difficulty recalling the order of autobiographical events (Pathman, Doydum, & Bauer, 2013), supporting the concern that adult ratings of when a word is learned may be inaccurate. Given these concerns regarding AoA ratings, researchers have attempted to validate these ratings. For example, Morrison, Chappell, and Ellis (1997) had children between the ages of 2 to 7 name pictures and, on the basis of their performance, estimated objective AoAs for the picture names. Importantly, the correlation between these objective AoA norms and adult ratings was high ($r = .747$).

Another method to combat the potential inaccuracy of adult ratings of AoA has been to use a frequency-based measure of when a word is first learned, also referred to as frequency trajectory. This variable was first introduced by Zevin and Seidenberg (2002) through their computational modeling efforts, which will be discussed in more detail in the next section. Briefly, frequency trajectory examines the changing pattern of frequency of words across childhood and therefore has been argued to be a more objective measure of when a word is first learned. If a word is used frequently in books read in early childhood, one could reason that it was learned at this stage. Conversely, if a word is not used in early childhood but appears with a greater frequency in high school or beyond, then it can be considered to be a late acquired word.

Zevin and Seidenberg (2002) introduced the concept of both frequency trajectory and cumulative frequency using the Educator's Word Frequency Guide (WFG; Zeno et al., 1995). The WFG, which is based on over 17 million word tokens, samples written textbooks, popular fiction, and nonfiction novels. Balota et al. (2004) found that the WFG explained the most variance in LDTs to over 2,400 monosyllabic English words (see also Brysbaert & New, 2009). The corpus of the WFG specifically contains writing samples relevant for beginning readers as well as skilled adult readers. In the corpus, books were assigned to different grade levels, 1 through 13 (i.e., college level). Some words are of a relatively high frequency in early grades, but become lower in frequency in adulthood (e.g., *dragon*) and therefore represent a high-to-low (HL) frequency trajectory. Other words show the opposite low-to-high (LH) trajectory (e.g., *oxygen*). *Cumulative frequency* refers to the frequency that a word is experienced throughout the life span and can be calculated by summing the grade-level frequencies. Zevin and Seidenberg (2002) argued that examining these trajectories (HL vs. LH) would give a good indication as to whether there is a true AoA effect in word recognition, because cumulative frequency of the words would be equated while varying the grade at which the word is exposed.

It is important to note that the underlying construct of interest for both adult AoA ratings and frequency trajectory, when used as a proxy for AoA, is the age at which a child first learns a word. Both variables have limitations associated with them as pure measures of word learning age and, as we discuss below, it has been argued that frequency trajectory is not as successful as rated AoA at measuring the objective age at which a word is first learned (Brysbaert, 2017). However, frequency trajectory also provides another interesting aspect of word experience that is not captured by rated AoA; a measurement of the pattern with which a given word occurs in books across an individual's education. Given this, the focus of the current research is on examining whether the two variables can account for unique and nonoverlapping variance in word recognition for young adults. Several studies have been conducted in a variety of tasks exploring the role of frequency trajectory. Prior to discussing these studies, we will first discuss predictions from computational models of AoA and frequency trajectory.

Computational Models of AoA and Frequency Trajectory

Ellis and Lambon Ralph (2000) were the first to explore the influence of AoA in a three-level connectionist model. Items were entered into training at different points. Over the course of several simulations, they demonstrated that items entered into training early were learned more effectively by the system, as long as the early learned items continued to be presented throughout the training period. Interestingly, this effect persisted even if the early learned items were significantly reduced in frequency of occurrence in later training. These modeling efforts led to the formulation of a possible locus of AoA effects, the network plasticity hypothesis. According to this hypothesis, early learned words gain an advantage in processing due to the fact that the network has greater plasticity earlier in training. As items are added to the network later in training, there is less plasticity and thus these items are not learned as effectively (see also Lambon Ralph & Ehsan, 2006; Smith, Cottrell, & Anderson, 2001).

Turning to Zevin and Seidenberg's (2002) computational modeling work, they questioned prior research on AoA, and instead suggested that reported AoA effects spuriously reflected uncontrolled differences in word frequency. To test this, a series of connectionist model simulations were conducted to explore the roles of both frequency trajectory and cumulative frequency on word reading performance of the models, with frequency trajectory serving as a proxy for AoA. These models differed in several respects from the Ellis and Lambon Ralph (2000) model, by including cleanup units and a large number of background items. The only simulation where an effect of frequency trajectory was observed was when the nature of the input-to-output mappings was arbitrary, a situation that does not reflect the learning of orthography-to-phonology relationships in English. They concluded that frequency trajectory, and therefore AoA, do not impact word naming in English. However, they did acknowledge that true AoA effects may be found in tasks where what is learned about early acquired items does not transfer to late acquired items. Examples include providing names for faces (e.g., Moore & Valentine, 1998) or generating semantic associates for words (e.g., Brysbaert, Van Wijnendaele, & De Deyne, 2000).

Using the same model architecture as Ellis and Lambon Ralph (2000); Mermillod, Bonin, Méot, Ferrand, and Paindavoine (2012) further explored the role of frequency trajectory by including more frequency trajectory conditions than were explored by Zevin and Seidenberg (2002). For three simulations (1, 2a, and 2b), they included late items, which were initially trained at a very low frequency and increased to high frequency at the end of training; early items, which were presented for a high frequency at the beginning of training but then decreased to a low frequency by the end; stable items, which were presented at a constant midrange frequency throughout; increasing frequency trajectory items, which were presented at a steady increasing frequency throughout training (i.e., LH items); and decreasing frequency trajectory items, which were presented at a steady decreasing frequency throughout training (i.e., HL items). They noted that these additional trajectories (especially the stable condition) more closely approximate the learning conditions of children. When the input-to-output mapping was arbitrary, they observed that the model performed best for both the early and decreasing HL sets, which did not significantly differ from each other. Stable items outperformed increasing LH items, which further outperformed late items. They interpreted this pattern as supporting the role of frequency trajectory for tasks with arbitrary mappings (e.g., picture naming). In two other simulations, the mappings between input-and-output were more systematic. Although they saw some influence of frequency trajectory at the end of training in both cases, they argued that these effects may be too small to exert an influence in behavioral tasks.

The modeling work of Monaghan and Ellis (2010) is the most relevant to the present study, as they explored the impact of AoA, cumulative frequency, and frequency trajectory in the same model. In addition, they estimated these variables through the use of the WFG (Zeno et al., 1995), which is the frequency index employed currently. Monaghan and Ellis (2010) used a developmental connectionist model of word naming based on that used by Harm and Seidenberg (1999). The developmental model had 14 epochs of training, representing the 13 grade levels in the WFG, plus one additional "adult reading" level. A total of 6,229 words were entered into training based on when they first appeared in the WFG at a frequency that was set to a given threshold. For example, for an item to be entered into training in Epoch 1, it must have a frequency of at least 1,000 per million for Grade 1 in the WFG. For Epoch 2, a frequency cut off of 100 per million in Grade 2 frequency was required. By Epoch 12, a frequency cutoff of 1 per million for Grade 12 frequency was used. There were significant effects of AoA at the end of training in this model. Analyses demonstrated that this was due to the position of entry of the items into the training process, as opposed to properties of the items themselves (e.g., length, cumulative frequency), thus providing a strong basis for a true AoA effect in the model. They also examined the role of frequency trajectory by comparing Grade 13 to Grade 1 frequency. The direction of the significant frequency trajectory effect was such that words decreasing in frequency during training had greater error at the end of training relative to words that increased in frequency. Therefore, LH items outperformed HL items at the end of training. The AoA effect was still significant in the model including frequency trajectory. Monaghan and Ellis therefore concluded that frequency trajectory may have a small influence on word naming performance even after AoA is controlled.

To summarize the main results from these models, Ellis and Lambon Ralph (2000) illustrated a clear effect of AoA in a model with arbitrary mappings. With respect to frequency trajectory (although it was not labeled as such), they saw equivalent performance for items trained at a high frequency throughout performance (i.e., high-to-high [HH] trajectory) and those whose frequency decreased from high-to-low across training. Zevin and Seidenberg (2002) directly compared only HL and LH trajectories, which they contended is the true test for AoA effects (or "age-limited learning" effects). Effects did not emerge when the models had systematic mappings between input and output but did so when the mappings were arbitrary. Mermillod et al. (2012) further explored additional frequency trajectories. Consistent with Zevin and Seidenberg they observed a benefit for decreasing trajectories (HL) relative to increasing trajectories (LH) that was greatest in a model with arbitrary mappings. Finally, the model of Monaghan and Ellis (2010) suggests that once AoA is controlled for, there is actually a decrement in model performance when words have a HL trajectory compared to a LH trajectory.

Empirical Studies of Frequency Trajectory

Since the initial introduction of the frequency trajectory variable by Zevin and Seidenberg (2002), several studies have explored its impact on lexical processing. Most of these studies have used frequency trajectory as a measure of "age-limited learning" to serve as a proxy for AoA. Consistent with the interpretation of Zevin and Seidenberg (2002), the assumption has been that HL items should be processed faster than LH items if age-limited learning effects do exist in word recognition. For present purposes, we will focus on the literature examining frequency trajectory effects on lexical decision performance, as this is the primary task used in the current study. Table 1 summarizes the effects of frequency trajectory on lexical decision performance in 13 assessments from four different studies. Frequency trajectory has been operationalized as a continuous variable, typically by subtracting a standardized measure of child frequency from a standardized measure of adult/college-level frequency (Bonin et al., 2004; Brysbaert, 2017; Lété & Bonin, 2013). It has also been examined factorially by exploring the impact of four frequency trajectory

Table 1
Summary of Main Results from Published Lexical Decision Studies of Frequency Trajectory

Study	Conditions examined	Variables included/controlled	Nature of effect (s)
Bonin, Barry, Méot, and Chalard (2004)	FT Continuous	CF, IMG, IV, Fam, N, Len, BF	Positive effect of FT; decreasing trajectories = faster RTs[a]
Caza & Moscovitch (2005)	HH, HL, LH, LL	Syl, Len, Phon, BF	LH faster than HL[b]
	HH, HL, LH, LL	Syl, Len, Phon, BF	LH faster than HL[c]
	HH, HL, LH, LL	Syl, Len, Phon, BF	LH faster than HL[d]
	HH, HL, LH, LL	Syl, Len, Phon, BF	No significant difference between LH & HL[e]
Lété & Bonin (2013)	HH, HL, LH, LL	Syl, Len, Consistency	LH was faster than LL, HH faster than HL[f]
	HH, HL, LH, LL	Syl, Len, Consistency	LH was faster than LL[g]
	FT Continuous	CF, Len, N, Consistency	ns[h]
	FT Continuous	Not Reported	Negative effect of FT; increasing trajectories = Faster RTs[i]
Brysbaert (2017)	FT Continuous	Objective AoA, CF	ns
	FT Continuous	Objective AoA, CF	ns
	FT Continuous	Objective AoA, CF	Positive effect of FT; decreasing trajectories = faster RTs[j]
	FT Continuous	Objective AoA, CF	ns

Note. FT = frequency trajectory; HH = high to high; HL = high to low; LH = low to high; LL = low to low; CF = cumulative frequency; IMG = imageability; IV = image variability; Fam = conceptual familiarity; N = orthographic neighborhood size; Len = word length in letters; BF = bigram frequency; Phon = number of phonemes; AoA = age of acquisition.
[a] FT was only significant when no measure of AoA was included in the regression. [b] Older adult group (ages 67–78). [c] Younger adult group (ages 21–28). [d] Individuals with dementia (ages 53–84). [e] Highly educated older adult group (ages 64–78). [f] Contrast of HH and HL was significant by participants only. [g] Factorial re-analysis from the French Lexicon Project (Ferrand et al., 2010). [h] Multiple regression from the French Lexicon Project (Ferrand et al., 2010). [i] Multiple regression from a subset of very inconsistent words from the French Lexicon Project (Ferrand et al., 2010). [j] Living Word Vocabulary (Dale and O'Rourke, 1981) was used as the Objective AoA variable.

conditions (HH, HL, LH, and low-to-low [LL]); Caza & Moscovitch, 2005; Lété & Bonin, 2013).

As can be seen from the table, the results are mixed. Frequency trajectory was not a significant predictor of LDTs in five analyses. In two analyses, words with a decreasing frequency trajectory resulted in faster LDTs. Bonin et al. (2004) found this to be the case only when a measure of AoA was not included in the model (which they argue is appropriate because frequency trajectory is being used as a measure of AoA). Brysbaert (2017) only saw an impact in one out of four analyses, each with a different measure of objective AoA included. In comparison, six of the analyses demonstrated that words with an increasing frequency were responded to faster in lexical decision. It should be noted that Caza and Moscovitch (2005) measured frequency trajectory differently than other studies. They included "dated" (e.g., accessory) words that changed from high frequency in childhood to low frequency in adulthood and "contemporary" (e.g., abstract) words that showed the opposite trajectory. They assessed dated and contemporary words by examining two older frequency corpora, the Thorndike and Lorge (1944) corpus and the Kučera and Francis (1967) corpus. They also included popular (e.g., *adequate*) and rare (e.g., *adamant*) words, which had a higher or lower frequency in both corpora, respectively. Using the same set of stimuli, they tested three groups of older adults and one group of younger adults. For three of their participant groups, the contemporary words, which display a LH frequency trajectory, were responded to significantly faster than the dated words, which display a HL trajectory. However, there are potential methodological concerns with the Caza and Moscovitch (2005) study. For example, the older frequency corpora that were used are not ideal for providing frequency information relevant for current college students as they are out-of-date and based on too few items (see Brysbaert et al., 2011). In addition, the contemporary and dated words were not matched on cumulative frequency and the sample size was also quite low in each study (15 participants in each).

In Lété and Bonin (2013), words with HH and LL trajectories were included in addition to HL and LH trajectory words in French. The HL and LH words were not found to significantly differ from each other on any task. On the other hand, LL words elicited significantly longer processing times than the LH words in lexical decision, among other tasks. HH words received a processing advantage compared to HL words in word naming and lexical decision (by participants). Lété and Bonin (2013) also conducted a regression analysis of over 26,000 words from the French Lexicon Project (Ferrand et al., 2010). Frequency trajectory did not significantly predict LDTs in the overall analysis. Interestingly, when they restricted their analysis of the French Lexicon Project to only "very inconsistent words," they reported a significant effect of frequency trajectory that was in the opposite direction of what they predicted. Words that became *higher* frequency in adulthood were processed faster than words that were higher frequency in childhood relative to adulthood (Footnote 7, Lété & Bonin, 2013). The authors suggest this pattern may be due to strategic factors related to processing highly inconsistent words. However, this pattern was also observed by Caza and Moscovitch (2005) in three of their participant groups.

In contrast to the mixed results with lexical decision, the studies exploring frequency trajectory effects in word naming have consistently demonstrated no influence of this variable (Bonin et al., 2004; Cuetos & Barbón, 2006; Zevin & Seidenberg, 2004). The only exception is reported by You, Chen, and Dunlap (2009) in Chinese and was restricted to cases where the characters did not have a phonetic radical, thereby decreasing the predictability of their pronunciation. In this case, HL characters were named significantly faster than the LH characters. This is consistent with the models of Zevin and Seidenberg (2002) and Mermillod et al. (2012), because the relationship between input and output was arbitrary. Frequency trajectory has been found to impact picture naming, a task that taps into semantic representations (Bonin et al., 2004; Bonin, Méot, Mermillod, Ferrand, & Barry, 2009); however, this is only the case when a measure of objective AoA is not included in the regression model (e.g., Pérez, 2007).

Frequency Trajectory as an Index of Word Experience

As discussed above, much of the interest in the frequency trajectory variable has been related to questions surrounding the measurement of AoA, with some researchers arguing that frequency trajectory is a better measure of when words are first learned than adult ratings or performance-based objective measures such as children's ability to name pictures (e.g., Bonin et al., 2004, 2009; Lété & Bonin, 2013; Mermillod et al., 2012; Zevin & Seidenberg, 2002, 2004). In support of this, several studies have found that frequency trajectory is a significant predictor of adult ratings of AoA (e.g., Bonin et al., 2004; You et al., 2009; Zevin & Seidenberg, 2004) as well as objective AoA (Bonin et al., 2004). However, Brysbaert (2017) examined the criterion validity of frequency trajectory compared to AoA ratings. He noted that the correlation between rated AoA and objective AoA was much higher than between frequency trajectory and objective AoA. Regression analyses for objective measures of AoA showed significant effects of rated AoA but no influence of frequency trajectory when both were included with other variables. Thus, although frequency trajectory is significantly correlated with AoA ratings (Bonin et al., 2004, 2009; Brysbaert, 2017), it does not seem to represent an adequate objective measure of when a word was acquired.

Setting the debate regarding the measurement of AoA aside, exploring the impact of frequency of usage across childhood can provide insight into how experience with words at certain points in life impacts current word recognition processes. Does the frequency of words in early childhood still impact performance in college? Alternatively, is the current frequency of the word as experienced in college the more important factor? The nature of the WFG corpus (Zeno et al., 1995) makes it possible to examine whether the patterns of frequency exposure over a college student's previous education influences the current processing of words. There has been some suggestion that current frequency of exposure may be more important than early frequency. For example, Joseph, Nation, and Liversedge (2013) recorded eye movements of children (mean age = 8.4 years) and adults reading sentences that contained a target word that was either high or low in child frequency. The word frequency manipulation was based on child frequency corpora and were controlled on adult frequency as well as rated AoA. Child frequency had an earlier impact on the eye movement record for the children than for adults. Gaze durations, which sum fixations on the word the first time it is read, were significantly longer for low frequency words relative to high frequency words for children. Adults only showed an effect on a later measure, total reading times, which takes refixations into account. Thus, the influence of child frequency was more immediate for children compared to adults. This could indicate that current frequency of exposure is the more important variable.

The Present Study

Based on extant research, the frequency trajectory literature remains contentious. The reliability and even direction of the frequency trajectory effect are unclear, and there is also debate on whether frequency trajectory faithfully reflects the age at which a child first learns a word (Brysbaert, 2017). The overarching purpose of the present study is to examine the impact of frequency trajectory in college students in a rigorous and systematic manner, so as to better assess the relative contributions of childhood frequency and current frequency in word processing times. To address this question, we conducted a new lexical-decision task with college students. In line with Lété and Bonin (2013), we examined words with "flat" trajectories (HH and LL) as well as words with decreasing (HL) and increasing (LH) trajectories. This allows an examination of the impact of early childhood exposure and current exposure, and if these factors interact.

To corroborate the results of the factorial study, item-level regression analyses ($N = 3{,}039$) were next conducted to examine the impact of frequency trajectory on LDTs and naming times from the English Lexicon Project (ELP; Balota et al., 2007). Compared to previous factorial studies in this domain, we controlled for many more word properties, and also incorporated a novel operationalization of the frequency trajectory variable using the WFG corpus. Importantly, the models included AoA and cumulative frequency, affording a further exploration of these variables' effects on word processing. Finally, to establish the robustness of the analyses on the ELP, parallel analyses were conducted on a second lexical decision dataset ($N = 2680$) drawn from the BLP (Keuleers et al., 2012).

Experiment 1: Lexical Decision Experiment

Method

Participants. Forty-four undergraduates at Wesleyan University (20 women, 24 men) received partial course credit for their introductory psychology course for participation. All participants were over the age of 18 and were enrolled in Wesleyan as undergraduates. Out of the 44, a total of 26 students were in their first year of college, 13 students were in their second year, four students were in their third year, and one student was in his or her fourth year. A prescreening indicated that English was not the primary language of two of the participants, so they were removed from analysis. Approval for recruiting participants was granted from the Wesleyan University Psychology Department Ethics Committee.

Materials. A total of 96 words were selected from the Educator's WFG (Zeno et al., 1995) to fulfill the four frequency trajectory conditions in the present study; HH, HL, LH, and LL. Twenty-four words were selected per condition. Example stimuli and frequency trajectories are displayed in Table 2. The WFG provides what is referred to as a "U" frequency for each grade level, which represents per million frequency weighted by a measure of dispersion across different content areas contained within the corpus. In the corpus, a value of zero is given to grade levels where the U value is less than 1 per million and a blank value is given for grade-levels where the word was not observed. Computations of frequency trajectories in Experiment 1 were based on the frequencies in Grade 1 and Grade 13 (i.e., college level). Words in

Table 2
Example Items from the Four Frequency Trajectory Conditions in Experiment 1

Condition	1	2	3	4	5	6	7	8	9	10	11	12	13+
World	150	273	399	476	663	720	738	745	750	777	766	789	766
Uncle	233	221	180	155	122	99	84	79	64	47	38	25	6
Brain	3	25	42	42	55	64	64	63	70	80	77	88	143
Opera	0	0	3	3	4	5	6	6	7	5	4	4	4

Note. Grade level frequencies were taken from the Educator's Word Frequency Guide (Zeno, Ivens, Hillard, & Duvvuri, 1995).

the high frequency conditions had a frequency of over 70 per million or higher in the relevant grade level and all words in the low frequency conditions had a frequency of less than 20 per million. A blank value in the grade-level corpora was considered to be a 0 value for that grade level. Cumulative frequencies were calculated by summing the per million frequency estimates across all of the grade levels. The items were rated for AoA (using the Gilhooly & Logie [1980] 1–7 scale) by Wesleyan undergraduates, where higher numbers indicate a later age of acquisition. They were included with additional items and split into two questionnaires. A total of 10 undergraduates rated each questionnaire. These ratings correlated strongly with ratings collected by Kuperman, Stadthagen-Gonzalez, and Brysbaert (2012), $r = .857$, $p < .001$, for the 95 overlapping item. AoA ratings varied by condition. Table 3 presents the item means for each condition.

Procedure. The 96 target items were combined with 32 filler words that were all five to seven letters in length. In addition, 127 nonwords were created by changing one or two letters in an existing five- to seven-letter English word to make a pronounceable and orthographically legal nonword. Stimuli were presented in the center of a computer screen. The font was Courier New 18 pt and was presented in black (in boldface) in lower case on a white background. E-Prime 2.0 (Psychology Software Tools, Inc. Sharpsburg, PA) was used to display the stimuli and record the responses on a standard keyboard. The "M" key was labeled with a "W" to indicate a word response and the "Z" key was labeled with an "NW" to indicate a nonword response. All participants were tested individually in a quiet testing room. They were instructed to sit at a comfortable distance from the computer keyboard and monitor.

Instructions were given on the computer screen at the beginning of the experimental session. This was followed by a practice block consisting of 15 items (eight words and seven nonwords). Prior to the display of each item, a fixation cross was presented in the center of the screen for 1,000 ms. The item was then displayed until a response was made or 10,000 ms had elapsed. After the practice block, participants were given a short break. They then completed 240 trials, of which 96 were the critical target items.

Data analysis. The data analysis was conducted using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) within the R environment for statistical computing (Version 3.2.1; R-Core Development Team, 2015). A linear mixed effects regression model (LMM) was conducted on the LDTs. Grade 1 and Grade 13 frequencies were added to the models using a −.5 coding for the high frequency condition and a .5 coding for a low frequency condition. The model including random slopes for participants failed to converge, so the simpler model including only random intercepts for both participants and items was adopted. Examining the main effects allow a test of the impact of Grade 1 frequency on its own and college frequency on its own when the respective other frequency is controlled. A significant interaction indicates that the pattern of frequency usage (i.e., the trajectory) impacts processing.

Table 3
Descriptive Statistics for the Four Frequency Trajectory Conditions in Experiment 1

Condition	Grade 1 frequency	Grade 13 frequency	Cumulative frequency	Length	Rated AoA
High-to-high	189.83 (102.0)	219.25 (164.4)	3052.13 (1691.2)	5.46 (.6)	2.55 (.7)
High-to-low	162.04 (105.7)	8.25 (5.2)	754.71 (380.1)	5.42 (.6)	2.22 (.4)
Low-to-high	4.79 (4.4)	149.04 (92.6)	793.17 (387.4)	5.54 (.6)	3.80 (.8)
Low-to-low	4.54 (4.1)	3.17 (2.6)	94.54 (60.1)	5.54 (.6)	3.24 (.9)

Note. All frequencies were taken from the Educator's Word Frequency Guide (Zeno et al., 1995). Grade 1 and Grade 13 frequencies are per million. Cumulative frequency was calculated by summing the per million frequencies across all grades (1–13). Length is in number of letters. Age-of-acquisition (AoA) was rated on a 1–7 scale. Numbers in parentheses are the standard deviations. One-way analyses of variance were conducted on each characteristic: Grade 1 frequency, $F_2(3,92) = 43.996$, $p < .001$; college frequency, $F_2(3,92) = 30.825$, $p < .001$; cumulative frequency, $F_2(3,92) = 80.873$, $p < .001$; length, $F < 1$; AoA, $F_2(3,92) = 23.621$, $p < .001$. As planned, follow-up t-tests demonstrated that Grade 1 frequency was matched across the relevant conditions (high-to-high [HH] vs. high-to-low [HL] and low-to-high [LH] vs. low-to-low [LL], both $|t|$s < 1). College frequency was not significantly different for the HH and LH groups ($p > .05$) but was significantly higher for the HL compared to LL group, $t_2(46) = 4.31$, $p < .001$, even though the average for both conditions was in the low frequency range (under 10 per million). Also, as expected, the HH was significantly higher in cumulative frequency compared to the other three groups (all ps $< .001$). Although the HL and LH groups did not differ significantly in cumulative frequency ($|t| < 1$), both were higher than the LL condition (both ps $< .001$). Finally, as expected, rated AoA varied by condition. Follow-up t-tests indicated that each of the conditions differed significantly from the others in rated AoA (all ps $< .05$).

A general linear mixed effects regression (GLMM) was conducted for lexical decision accuracy using the same structure as for LDTs. For the LDTs, we report the regression coefficient, standard error, and *t*-value. Results are considered to be statistically at an alpha level of .05 if $|t|$ is greater than 1.96. For accuracy, we report the regression coefficient, standard error, z value, and *p* value associated with each test.

Significant interactions were followed-up with direct comparisons of the HH to the HL condition (to assess the role of a decreasing frequency trajectory), the LH to the LL condition (to assess the role of an increasing frequency), and the HL to the LH condition. These comparisons were the ones made by Lété and Bonin (2013) in their multitask analysis in French. The third comparison has most often been used to argue for a role of age-limited learning in lexical processing. For the comparisons, random slopes for participants were added to the models. Only correct responses were considered for the LDT analysis. In addition, LDTs 2.5 standard deviations above and below the condition means were removed from analysis. This led to the removal of approximately 2.34% of the data.

Results

Participant means are presented in Table 4. The LMM on LDTs demonstrated a significant effect of both Grade 1 frequency, $b = 22.48$, $SE = 4.00$, $t = 5.63$, and Grade 13 frequency, $b = 16.39$, $SE = 4.00$, $t = 4.10$. These main effects were qualified by a significant interaction between the variables, $b = 22.69$, $SE = 7.99$, $t = 2.84$. Comparisons demonstrated that the differences between the HH and HL conditions, $b = 4.94$, $SE = 4.99$, $t = 0.99$, and the HL and LH conditions, $b = 6.19$, $SE = 4.91$, $t = 1.26$, were not significant. However, the difference between the LH and LL conditions was significant, $b = 27.79$, $SE = 6.29$, $t = 4.41$.

For accuracy, there was again a main effect of Grade 1 frequency, $b = -0.62$, $SE = 0.16$, $z = -3.78$, $p < .001$. However, the effect of Grade 13 frequency did not reach significance, $b = -0.17$, $SE = 0.16$, $z = -1.03$, $p > .1$. The interaction between these two variables was again significant, $b = -0.96$, $SE = 0.33$, $z = -2.93$, $p < .01$. The comparisons mirrored the LDT analyses in showing only a significant effect for the LH compared to LL condition (LH vs. LL, $b = -0.63$, $SE = 0.24$, $z = -2.61$, $p < .01$; HH vs. HL, $b = 0.29$, $SE = 0.42$, $z = 0.70$, $p > .1$; HL vs. LH $b = -0.68$, $SE = 0.37$, $z = -1.80$, $p > .05$).

In a final set of analyses we added the local AoA ratings as a continuous covariate, centered on its mean, to the models. This allows an examination of the impact of frequency trajectory once rated AoA is controlled. For LDTs, adding AoA to the model did not change the pattern of significant effects (all *t*s > 3.00). For accuracy, adding AoA reduced the impact of Grade 1 frequency, which was now only marginally significant, $b = -0.38$, $SE = 0.20$, $z = -1.91$, $p = .057$. The effect of Grade 13 was still not significant ($p > .1$), although the interaction between variables remained significant ($p < .01$). Significant rated AoA effects were also evident in both models (LDT, $b = 6.06$, $SE = 2.81$, $t = 2.16$; accuracy, $b = -0.20$, $SE = 0.10$, $z = -2.06$, $p < .05$). In addition, the inclusion of rated AoA significantly improved the fit of both models (LDT, $\chi^2 = 4.76$, $p = .029$; accuracy, $\chi^2 = 3.99$, $p = .046$).

Table 4
Lexical Decision Times (LDTs) and Error Rates for Experiment 1

Condition	LDTs	Error rates
High-to-high	498 (52.5)	3.47 (4.7)
High-to-low	503 (54.9)	2.58 (3.7)
Low-to-high	509 (54.9)	3.97 (4.1)
Low-to-low	536 (54.9)	7.24 (6.4)

Note. Means are computed by participants. LDTs are in milliseconds. Numbers in parentheses are the standard deviations.

Discussion

Overall, the experiment demonstrated an influence of both childhood (Grade 1) and current (college) frequencies on lexical decision performance. Words that were experienced with a high frequency in Grade 1 received significantly shorter LDTs than words that were experienced with a lower frequency. This effect persisted even when rated AoA was added to the model. Similarly, words that are experienced with a high frequency in college received significantly shorter LDTs on average when compared to words that are lower frequency in college texts. However, these main effects were qualified by an interaction indicating that when a word is experienced with a high frequency in childhood, current frequency does not impact lexical decision performance. The HH and HL conditions did not differ significantly in either LDTs or error rates even though both college frequency and cumulative frequency was significantly higher in the HH compared to the HL condition. This pattern of results differs from that observed in French by Lété and Bonin (2013). They observed that the HH condition was faster than the HL condition in lexical decision, although this effect only reached significance by participants in their analyses.

On the other hand, when a word is experienced with a low frequency in childhood, the current frequency of exposure has a large influence on lexical decision performance. The LH condition was responded to significantly faster and more accurately than the LL condition. In fact, the LL condition appeared to be quite slow and error prone overall suggesting that words that remain low frequency throughout childhood are quite difficult to discriminate from nonwords. This pattern of results is similar to that reported by Lété and Bonin (2013) in French, who observed that the LL condition was significantly slower than the LH condition. The interaction between Grade 1 and college frequency was also reliable when rated AoA was included in the analyses as a covariate. This shows that the interactive effects cannot simply be a function of a word's age of acquisition, which varied between the conditions. Importantly, including rated AoA in the models significantly improved their fit. This demonstrates that BOTH rated AoA and pattern of exposure over schooling impact lexical decision performance in college students. We will further discuss the joint impact of these variables on word recognition in the General Discussion.

Experiment 2: ELP Analysis

The results from Experiment 1 indicate that frequency trajectory does play a role in word recognition. The interaction between child frequency and college frequency suggests that when a word is encountered early in life with a high frequency, the current frequency of exposure has a minimal influence on processing. On the other hand, when the word is initially only encountered with a low frequency in childhood, the influence of current frequency is large and robust. Importantly, these effects were still observed when rated AoA was entered as a covariate, suggesting that both when a word is first encountered (as measured by rated AoA) and the pattern of exposure throughout childhood and adolescence (as measured by frequency trajectory) influence lexical processing. However, Experiment 1 used a factorial approach to examine frequency trajectory. Although 96 words (24 per condition) is reasonable for such a design, word recognition research is currently being advanced through the use of a "mega-study" approach (see Balota, Yap, Hutchinson, & Cortese, 2012). Exploring the impact of a variable of interest using a megastudy approach can provide further insight into the variable's contribution to word recognition processes for a larger, more representative, set of words.

Using the megastudy approach also allows us to examine frequency trajectory as a continuous measure, which has been accomplished in various ways in past studies, as discussed in the introduction. The WFG corpus provides information on Grades 1–13. It is therefore possible to examine the trajectory across all grade levels, as opposed to only focusing on the extremes, as was the case for Experiment 1 as well as past studies using this corpus to examine frequency trajectory (Brysbaert, 2017; Zevin & Seidenberg, 2002). To accomplish this, a log transform was first applied to each grade level frequency to reduce the skew of the distribution and a regression analysis was computed for each item. This yields a regression coefficient or slope for each item that represents the item's frequency trajectory. A negative slope indicates that the word decreased in frequency across the grade levels, a positive slope indicates that the word increased in frequency across grade levels, and a slope near zero indicates a flat trajectory. Figure 1 provides a boxplot illustrating the distribution of the

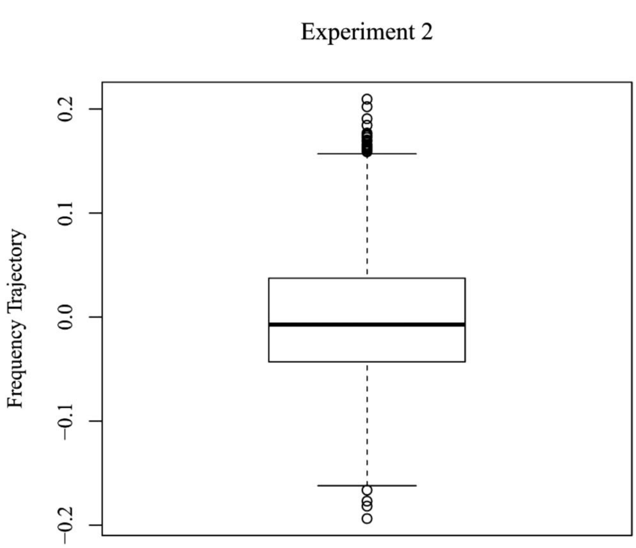

Figure 1. Boxplot illustrating the distribution of frequency trajectories for the 3,039 words used as stimuli in Experiment 2.

frequency trajectory variable for the items used in this analysis. We also examined the impact of cumulative frequency. Zero-order correlations between these variables and z-score standardized lexical decision latencies from the ELP (Balota et al., 2007) were assessed.

Using the megastudy approach also allows the influence of additional variables to be considered simultaneously. Importantly, both the influence of AoA and of frequency trajectory can be simultaneously assessed in the model. In the present analyses, frequency trajectory was added to the regression after the variability due to word onsets and 11 variables related to lexical and semantic processing were already accounted for. As illustrated in Table 1, the present analyses include many more variables than have been controlled for in past studies investigating frequency trajectory. These analyses therefore provide a stringent test of whether frequency trajectory has a reliable influence on lexical processing over and above other variables that have been shown to affect lexical decision performance. Finally, we also explored the influence of these variables on word naming latencies to compare the contribution of frequency trajectory across tasks.

Method

To evaluate the impact of the continuous measure of frequency trajectory on word recognition performance, the standardized RTs were extracted from the ELP (Balota et al., 2007) for 3,039 items for which all relevant predictors were available. A hierarchical by-items regression analysis was then computed in four steps. Below we outline the variables that were added to each step.

- Step 1: Following the work of Balota et al. (2004), the first step contained 13 dichotomous variables that coded for the characteristics of initial phonemes within the word. There was no variance for two of the onset predictors, so they were not included in the model.
- Step 2: This step included additional predictors that are related to lexical processing including word length, orthographic neighborhood size, phonological neighborhood size, as well as the mean orthographic and phonological Levenshtein distance to the 20 closest word neighbors (Orthographic Levenshtein Distance-20 [OLD-20] and Phonological Levenshtein Distance-20 [PLD-20]; Yap & Balota, 2009; Yarkoni, Balota, & Yap, 2008). In addition, the number of syllables and number of morphemes were included as was a contextual distinctiveness measure from the SUBTLEX-US corpus (Brysbaert & New, 2009). A measure of rated familiarity (Nusbaum, Pisoni, & Davis, 1984) was also included in this step.
- Step 3: This step included AoA ratings (Kuperman et al., 2012) as well as ratings of concreteness values (Brysbaert, Warriner, & Kuperman, 2014). We refer to this step as semantic variables, although we do note that there is debate as to whether the influence of AoA has a semantic locus or whether its influence is lexical in nature (see Juhasz, 2005, for a review of theories of AoA).
- Step 4: The frequency measures derived from the Zeno et al. (1995) WFG corpus were included in Step 4. For each item, a log transform was applied to the grade level frequency. In addition, the log transformed frequency values for each grade were summed to assess the impact of

cumulative frequency. A regression line was created for each item and the slope of the regression line was used as a measure of frequency trajectory.

Results

Table 5 displays the correlations of each of the predictors included in Steps 3–4 with the standardized LDTs and naming times from the ELP. Significant negative correlations were observed between the standardized RTs and cumulative frequency. The correlations of AoA and frequency trajectory with LDTs and naming latencies were positive and significant. For frequency trajectory, the magnitudes of both correlations were small, but statistically significant.

Table 6 displays the standardized regression coefficients for the predictor variables from Steps 2–4 for the standardized LDTs and naming times extracted from the ELP. Importantly, for both lexical decision and naming times, frequency trajectory was a significant predictor when added in Step 4 of the model. However, in comparison to the zero-order correlations, the direction of the effect is negative. This indicates that words that become higher frequency in later grades elicit shorter recognition times. Conversely, words that become lower frequency in later grades elicit longer recognition times. Interestingly, the effects of cumulative frequency were significant and inhibitory in both tasks, that is, words with higher cumulative frequency took longer to recognize. This pattern has been observed elsewhere in the literature when contextual diversity and word frequency are entered in the same regression model. For example, Adelman et al. (2006) reported that when contextual diversity was controlled for, word frequency produced either no unique effect or a suppression effect, in which high-frequency words yield slower responses. Importantly, the effect of rated AoA was also statistically significant in both tasks, indicating that words that are rated as being acquired earlier in life are generally responded to faster in visual word recognition.

Discussion

The results of the hierarchical regression analyses indicate that frequency trajectory is a significant predictor of LDTs for over 3,000 words when many other relevant predictor variables are accounted for in the regression equation. In contrast to past research (Bonin et al., 2004; Cuetos & Barbón, 2006; Zevin & Seidenberg, 2004), frequency trajectory was also found to be a significant predictor of word naming latencies. Importantly, the direction of the frequency trajectory effect on response times in the regression analyses is negative in both tasks, suggesting that words that increase in frequency over the course of education are responded to faster relative to words that decrease in frequency. This relationship is consistent with the pattern observed in Experiment 1 for the LH and LL conditions.

The negative relationship between frequency trajectory and response times differs from some reports of the effect of this variable in the literature. Some previous studies have reported that words that are higher in frequency in childhood relative to adulthood demonstrate a processing advantage in lexical decision (Bonin et al., 2004). However, Lété and Bonin (2013) failed to observe a significant effect of frequency trajectory for a large set of French words in lexical decision. As noted in the introduction, they did observe a significant negative effect of frequency trajectory when their analyses were restricted to highly irregular words. Although they dismiss this as a potential strategic effect, this pattern replicates that observed by Caza and Moscovitch (2005) in three of their participant groups in lexical decision with English words, as well as the current pattern. It also is consistent with the modeling work of Monaghan and Ellis (2010). It therefore stands to reason that the negative relationship of frequency trajectory with word recognition times may be a true reflection of how experience with words during education impacts their processing in adulthood.

It is also necessary to discuss the discrepancy in the direction of the frequency trajectory effect between the zero-order correlations (see Table 5) and the hierarchical regression analyses (see Table 6). When other variables are not taken into account, we see a small positive relationship of frequency trajectory with response times. It is only when frequency trajectory is included in a regression analyses with other relevant lexical and semantic variables does the negative relationship between this variable and lexical decision response times become clear. As a check to ensure that this negative relationship is not due only to the inclusion of cumulative frequency in Step 4, which is correlated with frequency trajectory, we also conducted a supplementary analysis where frequency trajectory was added as the sole predictor in Step 4. The regression coefficient for frequency trajectory was reliably negative for both

Table 5
Correlations Between Standardized Lexical Decision and Word Naming Times from the English Lexicon Project (Balota et al., 2007) and Key Predictors of Experiment 2 (N = 3,039)

Predictor	1	2	3	4	5	6	7
1. LDT Z-score	—	.518***	.415***	.025	−.312***	.045*	−.457***
2. NMG Z-score		—	.298***	.029	−.155***	.122***	−.311***
3. AoA			—	−.129***	−.271***	.444***	−.467***
4. Concreteness				—	.013	−.355***	−.317***
5. Familiarity					—	.025	.192***
6. Log slope						—	.068***
7. Log cumulative frequencies							—

Note. LDT = lexical decision times; NMG = naming; AoA = age-of-acquisition. Frequencies were derived from the Educator's Word Frequency Guide (Zeno et al., 1995). AoA was taken from Kuperman, Stadthagen-Gonzalez, and Brysbaert (2012) and concreteness was taken from Brysbaert, Warriner, and Kuperman (2014).
* $p < .05$. *** $p < .001$.

Table 6
Standardized Regression Coefficients from the Hierarchical Regression Analyses (N = 3,039) on Lexical Decision Times and Naming Times from the English Lexicon Project (Balota et al., 2007)

Variable	NMG (ELP)	LDT (ELP)
Step 1: Onsets		
Adjusted R^2	.241	.012
Step 2: Lexical variables		
LgSUBTL-CD	−.22***	−.43***
Familiarity	−.12***	−.21***
Number of morphemes	−.04*	−.02
Number of syllables	.06*	.08**
Number of letters	.24***	.12***
Orthographic *N*	−.09***	.07*
Phonological *N*	.12***	.04†
OLD-20	.02	.19***
PLD-20	.12***	.06†
Adjusted R^2	.496	.464
Change in R^2	.255***	.452***
Step 3: Semantic variables		
Concreteness	−.05***	−.13***
AoA	.08***	.05**
Adjusted R^2	.504	.482
Change in R^2	.008***	.018***
Step 4: Frequency trajectory and cumulative frequency		
Frequency trajectory	−.05**	−.17***
Cumulative frequency	.07**	.08***
Adjusted R^2	.506	.498
Change in R^2	.002***	.016***

Note. NMG (ELP) = naming (English Lexicon Project); LDT (ELP) = lexical decision time (English Lexicon Project); LgSUBTL-CD = Log of Subtitle Contextual Diversity; OLD-20 = Orthographic Levenshtein Distance-20; PLD-20 = Orthographic Levenshtein Distance-20. All variables were extracted from the English Lexicon Project (Balota et al., 2007) with the exception of age-of-acquisition (AoA), which was taken from Kuperman et al. (2012), familiarity which was taken from Nusbaum, Pisoni, and Davis (1984), concreteness which was taken from Brysbaert et al. (2014), and the frequency variables, which were taken from the Educator's Word Frequency Guide (Zeno et al., 1995). Coefficients reflect those entered in the corresponding step.
† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

naming (β = −.036, p = .03) and lexical decision (β = −.153, $p < .001$) times.

Experiment 3: BLP Analysis

The results of Experiment 2 demonstrate that frequency trajectory influences LDTs and naming times for over 3,000 English words. Both analyses showed that when a word increases in frequency across childhood and adolescence, it is responded to faster in word recognition tasks. This runs counter to some proposals for this variable, which view frequency trajectory as an objective measure of when a word is first learned, and therefore suggest that words with a HL trajectory should be processed faster than words with a LH trajectory. Instead, the present analyses indicate that both when a word is learned (as measured by rated AoA) and its pattern of exposure (as measured by frequency trajectory) can influence processing of the word.

However, given that there are contradictory reports in the literature about the presence and direction of frequency trajectory effects (see Table 1), it is necessary to replicate the current findings to insure their generalizability. We therefore conducted another megastudy analysis with an independent database. The BLP provides LDTs for over 28,000 mono- and disyllabic words (Keuleers et al., 2012). A hierarchical regression analysis was performed on the standardized LDTs for the 2,680 words contained in the database for which all relevant predictor variables were available. Figure 2 displays the boxplot of frequency trajectories for this set of items.

Method

The regression analysis on the standardized LDTs from the BLP was conducted in the same manner as in Experiment 2. The same predictor variables were used and were entered in four steps. As before, Step 4 consisted of the frequency trajectory variable and cumulative frequency which were derived from the Zeno et al. (1995) corpus in the same manner.

Results and Discussion

The standardized regression coefficients for Steps 2–4 of the hierarchical regression are displayed in Table 7. This analysis also demonstrated a negative relationship between frequency trajectory and LDTs (β = −.16), such that words that become more frequent in later adolescence and college enjoyed a processing advantage. Consistent with the ELP analyses, there was also an inhibitory effect of cumulative frequency once all other variables have been controlled. This will be returned to in the General Discussion.

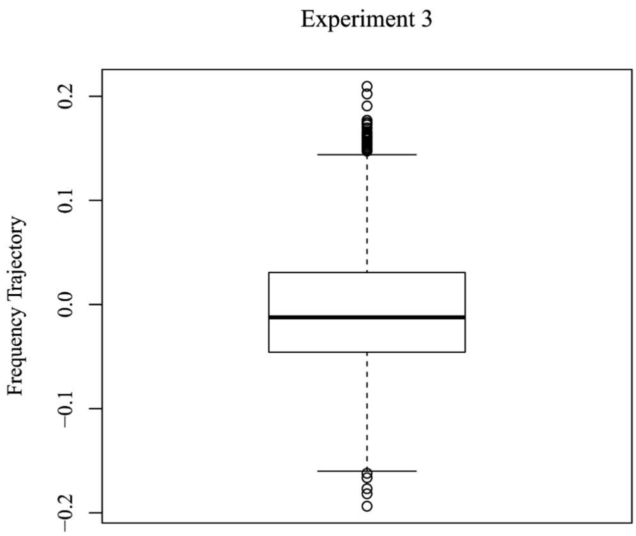

Figure 2. Boxplot illustrating the distribution of frequency trajectories for the 2,680 words used as stimuli in Experiment 3.

Finally, the effect of AoA was also apparent and in the expected direction; words that are rated as being acquired early in life received shorter LDTs. Thus, this analysis confirmed that both the age at which a word is learned and the pattern of its exposure during childhood and adolescence have significant and independent effects on lexical processing.

General Discussion

The present study examined whether the pattern of exposure to words in childhood impacts current word recognition processes for college students. Using the Zeno et al. (1995) WFG corpus, the frequency of words in both Grade 1 as well as in college can be assessed. Using a factorial approach, Experiment 1 suggested that for words that are initially experienced with a low frequency in Grade 1, the current frequency of occurrence in college has a robust effect on lexical decision performance. In contrast, for words initially experienced with a high frequency in Grade 1, college-level frequency had no effect. In the second experiment, we explored frequency trajectory as a continuous variable for a much larger number of items. We observed a negative relationship between frequency trajectory and both LDTs and word naming latencies to words from the ELP (Balota et al., 2007), when it was included in a last step of a hierarchical regression analysis after controlling for other relevant lexical and semantic variables. This analysis suggests that words that increase in frequency across early education are processed faster in adulthood. This finding was confirmed in Experiment 3 through the use of another independent large database of LDTs, the BLP (Keuleers et al., 2012). Below we examine whether the measurement of frequency trajectory in the present study impacted the results. This is followed by a discussion of the joint influence of AoA and frequency trajectory. Finally, we explore the role of frequency trajectory in light of extant theories and models of lexical processing.

The Measurement of Frequency Trajectory

The manner in which the continuous frequency trajectory measure was calculated in Experiments 2 and 3 differs from past research using this variable. To examine whether a different calculation would affect the pattern of results, frequency trajectory was recomputed using the procedure described in Brysbaert (2017), who adopted the approach of Zevin and Seidenberg (2002, 2004) and Bonin et al. (2004). Specifically, the Zeno et al. (1995) WFG frequencies for the earliest three grades were summed, log transformed, and then a z-score was calculated. The same process was repeated for the latest three grades in the corpus. Frequency trajectory was defined as the difference between the latest grades z-scores and the earliest grades z-scores. Cumulative frequency was defined as the sum of these z-scores.

The correlations between these new measures of frequency trajectory and cumulative frequency and those used in Experiments 2 and 3 were remarkably high (frequency trajectories $r = .978$, $p < .001$; cumulative frequencies $r = .993$, $p < .001$). Substituting these new measures into the regression models did not alter the nature of the frequency trajectory effect (ELP naming $\beta = -.040$, $p = .013$; ELP LDT $\beta = -.163$, $p < .001$; BLP LDT;

Table 7
Standardized Regression Coefficients from the Hierarchical Regression Analyses (N = 2,680) on Lexical Decision Times from the British Lexicon Project (Keuleers et al., 2012)

Variable	LDT (BLP)
Step 1: Onsets	
Adjusted R^2	.007
Step 2: Lexical variables	
LgSUBTL-CD	$-.49^{***}$
Familiarity	$-.25^{***}$
Number of morphemes	$.08^{***}$
Number of syllables	.01
Number of letters	.04
Orthographic *N*	−.01
Phonological *N*	$.06^{*}$
OLD-20	−.03
PLD-20	$.07^{*}$
Adjusted R^2	.399
Change in R^2	$.392^{***}$
Step 3: Semantic variables	
Concreteness	$-.11^{***}$
AoA	$.10^{***}$
Adjusted R^2	.421
Change in R^2	$.022^{***}$
Step 4: Frequency trajectory and cumulative frequency	
Frequency trajectory	$-.16^{***}$
Cumulative frequency	$.08^{**}$
Adjusted R^2	.436
Change in R^2	$.015^{***}$

Note. LDT (BLP) = lexical decision time (British Lexicon Project); LgSUBTL-CD = Log of Subtitle Contextual Diversity; OLD-20 = Orthographic Levenshtein Distance-20; PLD-20 = Phonological Levenshtein Distance-20. All variables were extracted from the English Lexicon Project (Balota et al., 2007) with the exception of age-of-acquisition (AoA), which was taken from Kuperman et al. (2012), familiarity which was taken from Nusbaum et al., (1984), concreteness which was taken from Brysbaert et al. (2014), and the frequency variables, which were taken from the Educator's Word Frequency Guide (Zeno et al., 1995). Coefficients reflect those entered in the corresponding step.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

$\beta = -.157$, $p < .001$). However, the inhibitory effect of cumulative frequency was no longer significant for the two LDT databases (ELP naming $\beta = .064$, $p = .007$; ELP LDT $\beta = .037$, $p = .123$; BLP LDT; $\beta = .045$, $p = .096$). These analyses thus support the finding that words with an increasing frequency trajectory are processed faster compared to words with a decreasing trajectory. Importantly, this pattern is not idiosyncratic to the manner by which we computed frequency trajectory.

The Unique Influences of AoA and Frequency Trajectory

It is worthwhile emphasizing that the effects of frequency trajectory observed in the present study are not inconsistent with findings that the age at which a word is learned is an important predictor of word recognition performance. Across all analyses, an effect of rated AoA was observed, demonstrating that words that are rated as being acquired earlier in life enjoy a processing advantage in adulthood. In Experiment 1, adding a measure of rated AoA to the linear mixed effects models significantly improved the fit of both the LDT and error rate analyses. Significant effects of AoA were observed in each regression model in Experiments 2 and 3. What the frequency trajectory findings do suggest is that once a word is learned, its pattern of frequency of exposure during childhood and young adulthood also impacts its processing.

As noted in the introduction, there has been debate over whether AoA and frequency trajectory should be included in the same model (e.g., Bonin et al., 2009). This debate has stemmed from the view that frequency trajectory can be considered to be an objective measure of when a word is initially learned and therefore may represent a better measure of age-limited learning than rated AoA. This view has recently been criticized by Brysbaert (2017). Hence, to explore whether the negative relationship between frequency trajectory and word processing time observed in the present study was due to the inclusion of AoA in the regression models in Experiments 2 and 3, these models were rerun with rated AoA removed. The coefficient for the frequency trajectory variable was negative in each instance although the ELP naming results were no longer significant (ELP naming, $\beta = -.004$, $p = .778$; ELP LDT $\beta = -.110$, $p < .001$; BLP LDT $\beta = -.087$, $p < .001$). In these analyses with AoA removed, the inhibitory effect of cumulative frequency was also no longer significant in the models (ps $> .09$). Thus, even when rated AoA was not included in the models for LDTs, words with an increasing frequency of occurrence across schooling resulted in faster LDTs in adulthood.

Taken together, the results of these experiments suggest that when a word is first acquired (as measured by rated AoA) and the pattern of exposure over the course of childhood and young adulthood collectively exert significant discernable effects on word processing in adulthood. On the other hand, the role of cumulative frequency was far more equivocal in the present study. In Experiment 1, the HH and HL conditions did not differ significantly from each other in LDTs or accuracy even though the cumulative frequency of the HH condition was more than four times greater than that of the HL condition on average (3052.13 vs. 754.71). Zero-order correlations with the ELP LDTs and word naming times reported in Table 5 showed moderate negative correlations with cumulative frequency. However, the impact of this variable was weakened when added in the fourth step of the regression analyses, with this cumulative frequency failing to reach significance in the majority of the re-analyses reported above.

The effect of cumulative frequency was also inhibitory in the regression analyses. This inhibitory effect may be due to the fact that a measure of contextual distinctiveness was included in our regression models. Work by Adelman and Brown (2008) has suggested that measures of the diversity of contexts in which a word appears is a more important variable to consider, compared to a simple count of how often a word occurs in text. Interestingly, Adelman et al. (2006) have demonstrated that when contextual diversity and word frequency were included in the same model, contextual diversity facilitated lexical decision and naming times, whereas word frequency was associated with either no effect or an inhibitory effect. In other words, our observation is nicely consistent with Adelman et al.'s (2006) findings.

Revisiting the Influence of Frequency Trajectory

If frequency trajectory is not simply a measure of when a word is first learned, why and how does this variable exert an impact on lexical processing? It may be useful to conceptualize this variable within the lexical quality hypothesis (Perfetti, 2007). According to this hypothesis, *lexical quality* is defined by precise and coordinated activation of a word's orthographic, phonological, and semantic representations. High lexical quality representations result in faster and more accurate word recognition. Lexical quality can vary between individuals (e.g., as a function of overall vocabulary size or spelling ability). In addition, lexical quality varies within a given individual's mental lexicon as a function of differences in the precision of word representations. The representations for a word may increase in their precision and accuracy as the word is experienced with a greater frequency over the course of a child's education. This would result in faster and more accurate lexical decision performance for these words relative to words that remain flat in their trajectories or decrease in frequency in later years.

It is estimated that at the beginning of kindergarten (ages 5–6), children have a working vocabulary of approximately 3,500 root words. This increases rapidly, and by the end of second grade, vocabulary is estimated to be approximately 6,000 root words (Biemiller, 2004). During early grades, children are beginning to learn how to read and are therefore developing orthographic representations for previously known words in support of reading. Experiencing the orthographic representation of a given word with a high frequency in Grade 1 may support the initial development of an accurate and well-specified orthographic representation that it tightly linked to the word's phonology and meaning. This will support word recognition into adulthood, as evidenced by the main effect of Grade 1 frequency on lexical decision performance in Experiment 1. However, in addition to this, a word that is experienced with an increasing frequency across childhood also results in faster processing in adulthood. The increasing frequency with which a word's orthographic representation is experienced may help to further develop the quality of that representation in the mental lexicon. Increasing frequency of viewing the word in text would also provide additional experience linking orthography to both phonology and meaning, which should result in a higher degree of lexical quality. This is not to say that frequency trajectory would be the only variable impacting lexical quality in adulthood. However, the increase in the pattern of exposure across schooling may be one mechanism that boosts lexical quality and therefore speeds lexical processing for certain words.

Although computational models of lexical decision differ in many respects (for a discussion, see Gomez, 2012), most include an important role for word frequency in making the word/nonword decision that is required for this task (e.g., Grainger & Jacobs, 1996; Ratcliff, Gomez, & McKoon, 2004). The current study expands on the importance of the word frequency variable by showing that the pattern of frequency across childhood impacts lexical decision performance. The impact of this variable is also supported by the computational modeling work of Monaghan and Ellis (2010), which was discussed in the introduction. Simulations on their developmental connectionist model of word naming revealed an effect of frequency trajectory, even after both AoA and cumulative frequency were controlled for. Specifically, words that increased in frequency yielded better naming performance at the end of training relative to words, which decreased in frequency. This pattern is of course consistent with our observation of frequency trajectory effects in both word naming and lexical decision, even when AoA is controlled.

Conclusion

In conclusion, the current results support the idea that both the age at which a word is learned as well as its pattern of exposure during childhood influences word recognition in adulthood. Words that are rated as being learned earlier in life are processed faster in both lexical decision and word naming tasks, supporting much past research on the importance of this variable. In addition, the factorial investigation supports a facilitatory role of Grade 1 frequency of occurrence on LDTs, indicating that the frequency with which a word is initially encountered in text (when learning to read) affects processing of that word in adulthood. At the same time, words that increase in their frequency across schooling also result in a processing advantage in adulthood. As discussed in the Introduction (and illustrated in Table 1), the evidence for frequency trajectory's effect on lexical processing has been equivocal in the literature. Both the factorial experiment and the multiple regression approach in the current study converge in showing that words that increase in frequency across schooling result in a processing advantage in adulthood. Therefore, these data provide empirical evidence for the existence of frequency trajectory effects in word recognition. It is our opinion based on these results that frequency trajectory should not be considered as a proxy for when a word is first acquired, thus supporting the main conclusion reached by Brysbaert (2017). However, the current studies do support a role for frequency trajectory in lexical processing, as a measure of word experience. These results extend the known importance of word frequency as a predictor of word recognition response times and suggest that patterns of exposure of words across childhood impacts the accuracy and precision of representations for words within the adult mental lexicon.

References

Adelman, J. S., Brown, G. D. A. (2008). Modeling lexical decision: The form of frequency and diversity effects. *Psychological Review, 115,* 214–229.

Adelman, J. S., Brown, G. D. A., & Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science, 17,* 814–823. http://dx.doi.org/10.1111/j.1467-9280.2006.01787.x

Baayen, R. H., Milin, P., & Ramscar, M. (2016). Frequency in lexical processing. *Aphasiology, 30,* 1174–1220. http://dx.doi.org/10.1080/02687038.2016.1147767

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1995). *The CELEX Lexical Database* [CD-ROM]. Philadelphia, PA: Linguistic Data Consortium.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133,* 283–316. http://dx.doi.org/10.1037/0096-3445.133.2.283

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., . . . Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39,* 445–459. http://dx.doi.org/10.3758/BF03193014

Balota, D. A., Yap, M. J., Hutchinson, K. A., & Cortese, M. J. (2012). Megastudies: What do millions (or so) of trials tell us about lexical processing? In J. S. Adelman (Eds.), *Visual word recognition Vol. 1: Models and methods* (pp. 90–115). Hove, East Sussex: Psychology Press.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67,* 1–48. http://dx.doi.org/10.18637/jss.v067.i01

Biemiller, A. (2004). Teaching vocabulary in the primary grades. In J. F. Baumann & E. J. Kame'enui (Eds.), *Vocabulary instruction* (pp. 28–40). New York, NY: Guilford Press.

Bonin, P., Barry, C., Méot, A., & Chalard, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story? *Journal of Memory and Language, 50,* 456–476. http://dx.doi.org/10.1016/j.jml.2004.02.001

Bonin, P., Méot, A., Mermillod, M., Ferrand, L., & Barry, C. (2009). The effects of age of acquisition and frequency trajectory on object naming: Comments on Pérez (2007). *The Quarterly Journal of Experimental Psychology, 62,* 1132–1140. http://dx.doi.org/10.1080/17470210802618850

Brysbaert, M. (2017). Age of acquisition ratings score better on criterion validity than frequency trajectory or ratings "corrected" for frequency. *The Quarterly Journal of Experimental Psychology, 70,* 1129–1139. http://dx.doi.org/10.1080/17470218.2016.1172097

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology, 58,* 412–424. http://dx.doi.org/10.1027/1618-3169/a000123

Brysbaert, M., & Cortese, M. J. (2011). Do the effects of subjective frequency and age of acquisition survive better word frequency norms? *The Quarterly Journal of Experimental Psychology, 64,* 545–559. http://dx.doi.org/10.1080/17470218.2010.503374

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, Instruments & Computers, 41,* 977–990. http://dx.doi.org/10.3758/BRM.41.4.977

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in Psychology, 7,* 1116. http://dx.doi.org/10.3389/fpsyg.2016.01116

Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica, 104,* 215–226. http://dx.doi.org/10.1016/S0001-6918(00)00021-4

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46,* 904–911. http://dx.doi.org/10.3758/s13428-013-0403-5

Burgess, C., & Livesay, K. (1998). The effect of corpus size in predicting reaction time in a basic word recognition task. Moving on from Kučera and Francis. *Behavior Research Methods, Instruments & Computers, 30,* 272–277. http://dx.doi.org/10.3758/BF03200655

Caza, N., & Moscovitch, M. (2005). Effects of cumulative frequency, but not frequency trajectory, in lexical decision times of older adults and patients with Alzheimer's disease. *Journal of Memory and Language, 53,* 456–471. http://dx.doi.org/10.1016/j.jml.2005.03.005

Cuetos, F., & Barbón, A. (2006). Word naming in Spanish. *European Journal of Cognitive Psychology, 18,* 415–436. http://dx.doi.org/10.1080/13594320500165896

Dale, E., & O'Rourke, J. (1981). *The living word vocabulary, the words we know: A national vocabulary inventory*. Chicago, IL: World Book.

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1103–1123. http://dx.doi.org/10.1037/0278-7393.26.5.1103

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., . . . Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods, 42,* 488–496. http://dx.doi.org/10.3758/BRM.42.2.488

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation, 12,* 395–427. http://dx.doi.org/10.3758/BF03201693

Gomez, P. (2012). Mathematical models of the lexical decision task. In J. Adelman (Ed.), *Visual word recognition* (pp. 70–89). Hove, UK: Psychology Press.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review, 103,* 518–565. http://dx.doi.org/10.1037/0033-295X.103.3.518

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review, 106,* 491–528. http://dx.doi.org/10.1037/0033-295X.106.3.491

Johnston, R. A., & Barry, C. (2006). Age of acquisition and lexical processing. *Visual Cognition, 13,* 789–845. http://dx.doi.org/10.1080/13506280544000066

Joseph, H. S. S. L., Nation, K., & Liversedge, S. P. (2013). Using eye movements to investigate word frequency effects in children's sentence reading. *School Psychology Review, 42,* 207–222.

Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin, 131,* 684–712. http://dx.doi.org/10.1037/0033-2909.131.5.684

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods, 44,* 287–304. http://dx.doi.org/10.3758/s13428-011-0118-4

Kučera, H., & Francis, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44,* 978–990. http://dx.doi.org/10.3758/s13428-012-0210-4

Lambon Ralph, M. A., & Ehsan, S. (2006). Age of acquisition effects depend on the mapping between representations and the frequency of occurrence. Empirical and computational evidence. *Visual Cognition, 13,* 248–292. http://dx.doi.org/10.1080/13506280544000110

Lété, B., & Bonin, P. (2013). Does frequency trajectory influence word identification? A cross-task comparison. *The Quarterly Journal of Experimental Psychology, 66,* 973–1000. http://dx.doi.org/10.1080/17470218.2012.723725

Mermillod, M., Bonin, P., Méot, A., Ferrand, L., & Paindavoine, M. (2012). Computational evidence that frequency trajectory theory does not oppose but emerges from age-of-acquisition theory. *Cognitive Science, 36,* 1499–1531. http://dx.doi.org/10.1111/j.1551-6709.2012.01266.x

Monaghan, P., & Ellis, A. W. (2010). Modeling reading development: Cumulative, incremental learning in a computational model of word naming. *Journal of Memory and Language, 63,* 506–525. http://dx.doi.org/10.1016/j.jml.2010.08.003

Moore, V., & Valentine, T. (1998). Naming faces: The effect of AoA on speed and accuracy of naming famous faces. *The Quarterly Journal of Experimental Psychology, 51*(A), 485–513.

Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology, 50*(A), 528–559.

Murray, W. S., & Forster, K. I. (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review, 113,* 327–357. http://dx.doi.org/10.1037/0033-295X.111.3.721

Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). *Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words* (Research on Speech Perception, Progress Report no. 10). Bloomington, IN: Speech Research Laboratory, Psychology Department, Indiana University.

Pathman, T., Doydum, A., & Bauer, P. J. (2013). Bringing order to life events: Memory for the temporal order of autobiographical events over an extended period in school-aged children and adults. *Journal of Experimental Child Psychology, 115,* 309–325. http://dx.doi.org/10.1016/j.jecp.2013.01.011

Pérez, M. A. (2007). Age of acquisition persists as the main factor in picture naming when cumulative word frequency and frequency trajectory are controlled. *The Quarterly Journal of Experimental Psychology, 60,* 32–42. http://dx.doi.org/10.1080/17470210600577423

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11,* 357–383. http://dx.doi.org/10.1080/10888430701530730

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review, 111,* 159–182. http://dx.doi.org/10.1037/0033-295X.111.1.159

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Smith, M. A., Cottrell, G. W., & Anderson, K. L. (2001). The early word catches the weights. In T. K. Leen, T. G. Dietterich, & V. Tresp (Eds.), *Advances in neural information systems* (pp. 52–57). Cambridge, MA: MIT Press.

Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York, NY: Teacher's College, Columbia University.

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language, 60,* 502–529. http://dx.doi.org/10.1016/j.jml.2009.02.001

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review, 15,* 971–979. http://dx.doi.org/10.3758/PBR.15.5.971

You, W., Chen, B., & Dunlap, S. (2009). Frequency trajectory effects in Chinese character recognition: Evidence for the arbitrary mapping hypothesis. *Cognition, 110,* 39–50. http://dx.doi.org/10.1016/j.cognition.2008.08.004

Zeno, S. M., Ivens, S. H., Hillard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Brewster, NJ: Touchstone Applied Science Associates.

Zevin, J. D., & Seidenberg, M. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language, 47,* 1–29. http://dx.doi.org/10.1006/jmla.2001.2834

Zevin, J. D., & Seidenberg, M. S. (2004). Age-of-acquisition effects in reading aloud: Tests of cumulative frequency and frequency trajectory. *Memory & Cognition, 32,* 31–38. http://dx.doi.org/10.3758/BF03195818

Received February 8, 2017
Revision received December 21, 2017
Accepted January 11, 2018 ■